Formulas for Cross Sections and Rate Coefficients for Partial Electron Impact Ionization of CH_4 and H_2

V. Dose,[a] P. Pecher, and R. Preuss

Max-Planck-Institut für Plasmaphysik, EURATOM Association, D-85748 Garching / München, Germany

Received 30 August, 1999; revised manuscript received 24 March 2000

Robust, accurate, and asymptotically exact cross section formulas for all direct and dissociative electron impact ionization channels of methane and hydrogen are presented. The parameter estimation employs Bayesian inference and allows for a consistent use of data from different experiments with possibly discordant calibrations. An efficient method for the calculation of rate coefficients and their temperature derivatives is outlined. © *2001 American Institute of Physics.* [S0047-2689(00)00205-1]

Key words: Bayesian probability theory; cross section; electron impact ionization; rate coefficient.

Contents

List of Tables

List of Figures

1. Introduction

Cross sections for electron impact ionization of methane and hydrogen are part of the necessary data base for modeling low temperature glow discharges. Interest in the latter stems from their application in the formation of diamond and diamondlike carbon thin films. A herculean effort also covering electron impact fragmentation of molecular ions and radicals as well as proton impact ionization of methane and methane derived radicals has been published by Erhardt and Langer.[1] The great need for such data condensed in suitable analytic form is reflected in the 35 citations of this paper up to today. Data on elementary processes in hydrogen–helium plasmas and corresponding analytic fits have been published by Janev *et al.*[2] We used both the methane and the hydrogen data in modeling mass spectra of the ion flux of a low temperature CH_4/H_2 discharge. In the course of the work it became clear that rate coefficients for electron temperatures as low as 2 eV were needed and this in turn caused a closer inspection of the experimental data and the Erhardt–Langer fit formula in particular in the threshold region. The situation is depicted in Fig. 1 where experimental data from two different sources are shown together with the fit as provided by Erhardt and Langer. Two disturbing features are revealed by this figure. First of all the two sets of experimental data are obviously not compatible with each other and second the Erhardt–Langer fit formula shows quite an unphysical bimodal cross section. The source of this peculiar behavior is very probably the fact that the fit function was a nine term polynomial in $\ln(x)$, where x is the electron impact energy in units of the threshold energy E_0, $x=E/E_0$. High order polynomial fits tend to such oscillations, as already pointed out by the authors themselves. For our needs the fit suffers moreover from a poor representation of the data in the threshold region. Both shortcomings become unimportant when rate coefficients for relatively large temperatures (say >50 eV) as in the boundary plasma of tokamaks are needed. For low temperature plasmas, however, they are unacceptable. The purpose of this paper is to show how to deal with the quite common situation where a model function shall be derived from data sets A and B, where A and B cover different regions and suffer from discording calibration. We shall use Bayesian inference, carry through the full Bayesian calculation, and develop approximate formulas which perform surprisingly well. *Partial electron impact ionization cross sections for CH_4 and H_2 are chosen to illustrate how the method works. However, we want to make quite clear at this point that we are not concerned with the critical evaluation of the measurement data itself.*

[a]Electronic mail: dose@ipp.mpg.de

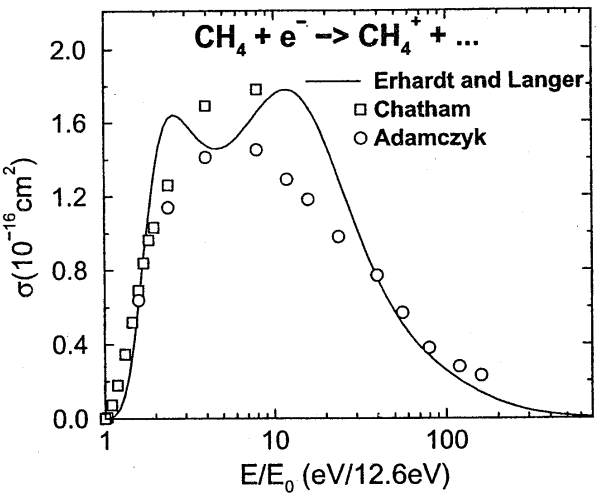

FIG. 1. Experimental cross sections for the electron impact ionization of methane. Note the discrepancy between the two data sets at $3 \lesssim E/E_0 \lesssim 8$. The line shows the polynomial fit by Erhardt and Langer (see Ref. 1).

2. Bayesian Rules

Bayesian probability theory (BPT) rests on the application of two rules. The first is the product rule which allows to expand a probability or a probability density function $P(H,D|I)$ of the two variables H and D conditional on further information I into simpler densities depending only on either H or D as a variable

$$P(H,D|I)=P(H|I)\cdot P(D|H,I)=P(D|I)\cdot P(H|D,I). \tag{1}$$

We identify H with hypothesis and D with data. Equating the two alternative expansions of $P(H,D|I)$ yields Bayes' theorem

$$P(H|D,I)=\frac{P(H|I)}{P(D|I)}\cdot P(D|H,I), \tag{2}$$

which tells us how to update prior knowledge $P(H|I)$ about hypothesis H in the light of data D collected from appropriate experiments designed to test H. $P(H|D,I)$ is called the posterior probability or probability density for H. It is the product of $P(H|I)$, the prior probability density for H, which contains all our knowledge about H prior to inspection of the data D. $P(H|I)$ may be entirely uninformative, it may reflect symmetry or invariance properties of H, or it may even encode prior knowledge about the numerical value of H and its possible range. $P(D|H,I)$ is the sampling distribution of the data when regarded as a function of D and is normalized as such or is called the likelihood if considered as a function of H. $P(D|H,I)$ is in other words the theory of the experiment which allows us to calculate the expected measurement data if we assume the physics (H) were known. The denominator $P(D|I)$ in Bayes' theorem is called the evidence or the prior predictive value or the global likelihood for the entire class of hypotheses characterized by H. It is not independent of the other probabilities. This follows from the second, the so-called marginalization rule of BPT

$$P(D|I)=\int P(H,D|I)dH=\int P(H|I)\cdot P(D|H,I)dH. \tag{3}$$

The denomination global likelihood for $P(D|I)$ is particularly suggestive on inspection of Eq. (3). $P(D|I)$ plays a crucial role in the comparison of different models given a data set D and the decision which of the given models is best represented by the data D. However, in the forthcoming context it serves merely as a constant which accounts for the correct normalization of the posterior probability for H. In this paper we shall apply Eqs. (1), (2), and (3) consistently and shall be content with these manipulations. We are well aware that this section is entirely inappropriate as an introduction to BPT, and refer the interested reader to excellent papers covering this subject in the literature.[3–6]

3. The Likelihood

Consider the case, where two sets $\vec{\delta}$ and $\vec{\Delta}$ of measured cross sections are available and were taken at energies $\vec{x}$ and $\vec{X}$. We then wish to relate the two sets of experimental data (the generalization to more sets is quite straight forward) to a single fit function. The general problem with this aim is that the two data sets cover usually different energy ranges and moreover may suffer from different calibrations. We assume that the fit function is fully specified by a linear parameter c establishing the scale of the fit function and a further set of (possibly) nonlinear parameters $\vec{\lambda}$ related to the shape of the fit function. The parametric dependence of the fit function on c and $\vec{\lambda}$ will become more transparent further down in Sec. 5. Let us further introduce an (unknown) calibration parameter γ which, when applied to the data set $\vec{\Delta}$, rescales it on the same scale as the set $\vec{\delta}$. All this information can be expressed by the probability density function

$$P(\vec{\delta},\vec{\Delta}|\vec{x},\vec{X},c,\gamma,\vec{\lambda},I)=P(\vec{\delta}|\vec{x},\vec{X},c,\gamma,\vec{\Delta},I)\cdot P(\vec{\Delta}|\vec{\delta},\vec{x},\vec{X},c,\gamma,\vec{\lambda},I). \tag{4}$$

The right-hand side of Eq. (4) follows from the product rule. The two factors on the right-hand side of Eq. (4) may be simplified. The conditional probability $P(\vec{\delta}|\vec{x},\vec{X},c,\gamma,\vec{\lambda},I)$ of course does not depend on the energies $\vec{X}$ at which the data $\vec{\Delta}$ are taken. Moreover, according to the convention of this paper the data $\vec{\delta}$ are considered to be correct on an absolute scale and consequently do not depend on γ which accounts for a calibration error in $\vec{\Delta}$. The above likelihood simplifies therefore to $P(\vec{\delta}|\vec{x},c,\vec{\lambda},I)$. By the same line of reasoning $P(\vec{\Delta}|\vec{\delta},\vec{x},\vec{X},c,\gamma,\vec{\lambda},I)$ turns out to be independent of $\vec{x}$, the set of energies at which the data $\vec{\delta}$ were taken and of $\vec{\delta}$ itself, the cross section data obtained in a different experiment.

This is true for any circumstances for which the rules of good scientific practice apply. Equation (4) simplifies therefore to

$$P(\vec{\delta},\vec{\Delta}|\vec{x},\vec{X},c,\gamma,\vec{\lambda},I)=P(\vec{\delta}|\vec{x},c,\vec{\lambda},I)\cdot P(\vec{\Delta}|\vec{X},c,\vec{\lambda},\gamma,I). \tag{5}$$

We now model the data by

$$\begin{aligned}\delta_i&=c\cdot\varphi(x_i,\vec{\lambda})+\alpha_i,\\ \gamma\Delta_j&=c\cdot\varphi(X_j,\vec{\lambda})+\beta_j.\end{aligned} \tag{6}$$

α_i and β_j are the respective errors. We assume that the expectation value of α_i, $\langle\alpha_i\rangle=0$ and $\langle\alpha_i^2\rangle=s_i^2$ and similarly for β_j, $\langle\beta_j^2\rangle=S_j^2$. These pairs of conditions do not, of course, specify corresponding distribution functions. In fact, an infinite manifold of functions can be found which meet the above moment conditions. By application of the principle of maximum information entropy[7] subject to the moment constraints the problem becomes unique and we obtain as the least informative distribution function meeting the moment constraints a Gaussian

$$P(\vec{\delta}|\vec{x},c,\vec{\lambda},I)=\prod_i \frac{1}{s_i\sqrt{2\pi}}\exp\left\{-\frac{1}{2}[\delta_i-c\varphi(x_i,\vec{\lambda})]^2/s_i^2\right\} \tag{7}$$

and

$$P(\vec{\Delta}|\vec{X},c,\vec{\lambda},\gamma,I)=\prod_j \frac{\gamma}{S_j\sqrt{2\pi}}\times\exp\left\{-\frac{1}{2}[\gamma\Delta_j-c\varphi(X_j,\vec{\lambda})]^2/S_j^2\right\} \tag{8}$$

The definitions

$$\begin{aligned}\varphi_i(x_i,\vec{\lambda})/s_i&=f_i \quad \varphi_j(X_j,\vec{\lambda})/S_j=F_j,\\ \delta_i/s_i&=d_i \quad \Delta_j/S_j=D_j\end{aligned} \tag{9}$$

allow us to introduce vector notation. The properly normalized likelihoods then read

$$P(\vec{d}|\vec{x},c,\vec{\lambda},I)=(2\pi)^{-n/2}\exp\left\{-\frac{1}{2}(\vec{d}-c\vec{f})^T(\vec{d}-c\vec{f})\right\}, \tag{10}$$

$$P(\vec{D}|\vec{X},c,\vec{\lambda},\gamma,I)=(2\pi)^{-N/2}\gamma^N\times\exp\left\{-\frac{1}{2}(\gamma\vec{D}-c\vec{F})^T(\gamma\vec{D}-c\vec{F})\right\}. \tag{11}$$

N is the number of data D_j and n the number of data d_i.

4. Parameter Estimation

Our first goal is the estimation of $\langle\mu^k\rangle$ from Eqs. (10) and (11), where μ stands for any of the parameters c, γ, $\vec{\lambda}$ and the brackets denote the expectation value. Of particular interest are of course $k=1,2$ from which we obtain the average value and the associated variance, respectively. In order to calculate $\langle\mu^k\rangle$ we need the probability density of the parameters given the data $(\vec{d},\vec{x},\vec{D},\vec{X})$. $P(c,\gamma,\vec{\lambda}|\vec{d},\vec{x},\vec{D},\vec{X},I)$ may be obtained by application of Bayes theorem Eq. (2)

$$P(c,\gamma,\vec{\lambda}|\vec{d},\vec{x},\vec{D},\vec{X},I)=\frac{P(c,\gamma,\vec{\lambda}|I)}{P(\vec{d},\vec{D}|\vec{x},\vec{X},I)}\cdot P(\vec{d},\vec{D}|\vec{x},\vec{X},c,\gamma,\vec{\lambda},I) \tag{12}$$

in terms of the previously specified likelihood function Eqs. (4), (10), and (11) and the prior density for the parameters $P(c,\gamma,\vec{\lambda}|I)$. The latter must be specified now. First of all, our knowledge about any of these parameters is independent of the value assigned to any other prior to taking the data of course. Consequently $P(c,\gamma,\vec{\lambda}|I)$ factorizes into

$$P(c,\gamma,\vec{\lambda}|I)=P(c|I)\cdot P(\gamma|I)\cdot P(\vec{\lambda}|I) \tag{13}$$

and no conditioning other than the general background information I enters the factors on the right-hand side due to mutual logical independence. Simple uninformative priors on c, γ, and $\vec{\lambda}$ are constants within a range specified by respective minimum and maximum values and zero outside. c and γ are further restricted to positive values. Occasionally we may have prior information on the amount by which the calibration in the experiment providing data $\vec{D}$ differs from that in the experiment providing data $\vec{d}$. Let g be the expectation value of γ, $\langle\gamma\rangle=g$, then by the principle of maximum entropy[7] $P(\gamma|g,I)=\exp(-\gamma/g)/g$. We shall employ this form of prior information on γ and can always return to the less informative situation letting $g\to\infty$ in the exponential. With this specification we can proceed and evaluate expectation values of a particular parameter. For reasons of definiteness we choose any of the components of $\vec{\lambda}$. Our first goal is then to obtain the marginal distribution $P(\vec{\lambda}|\vec{d},\vec{x},\vec{D},\vec{X},I)$ from Eq. (12) by application of the marginalization rule Eq. (3)

$$P(\vec{\lambda}|\vec{d},\vec{x},\vec{D},\vec{X},I)=\int d\gamma\int dc\exp\left\{-\frac{\gamma}{g}\right\}P(\vec{d},\vec{D}|\vec{x},\vec{X},c,\gamma,\vec{\lambda},I)/Z. \tag{14}$$

The normalization denominator Z summarizes the constant values of the prior in Eq. (13) and the evidence denominator in Eq. (12). The integrand in Eq. (14) is then given by Eqs. (4), (10), and (11) and hence

$$P(\vec{\lambda}|\vec{d},\vec{x},\vec{D},\vec{X},I)$$
$$=(2\pi)^{-(n+N)/2}/Z\int d\gamma\exp\left\{-\frac{\gamma}{g}\right\}\gamma^N\int dc$$
$$\times\exp\left\{-\frac{1}{2}(\vec{d}-c\vec{f})^T(\vec{d}-c\vec{f})\right.$$
$$\left.-\frac{1}{2}(\gamma\vec{D}-c\vec{F})^T(\gamma\vec{D}-c\vec{F})\right\}. \quad (15)$$

The argument of the exponentials in Eq. (15), E, may be comprised as

$$E=-\frac{A}{2}(c-c_0)^2-\frac{B}{2}. \quad (16)$$

Equating equal powers of c in the exponent of Eq. (15) with Eq. (16) we obtain

$$A=\vec{f}^T\vec{f}+\vec{F}^T\vec{F}, \quad (17a)$$

$$B=\vec{d}^T\vec{d}+\gamma^2\vec{D}^T\vec{D}-\frac{(\vec{d}^T\vec{f}+\gamma\vec{D}^T\vec{F})^2}{\vec{f}^T\vec{f}+\vec{F}^T\vec{F}}+2\frac{\gamma}{g}, \quad (17b)$$

$$c_0=(\vec{d}^T\vec{f}+\gamma\vec{D}^T\vec{F})/(\vec{f}^T\vec{f}+\vec{F}^T\vec{F}). \quad (17c)$$

The integral over c can then be done and results in general for $c_{\min}\leqslant c\leqslant c_{\max}$ in an expression involving two error functions. This complicated result simplifies if we assume that the numerical value of the integral changes only insignificantly if, in carrying out the integration, we replace the lower limit $c_{\min}$ by $-\infty$ and the upper limit $c_{\max}$ by $+\infty$. This assumption is well justified for sufficiently precise data $\vec{d}$, $\vec{D}$. The integral over c is then just $(2\pi/A)^{1/2}$ and we obtain

$$P(\vec{\lambda}|\vec{d},\vec{x},\vec{D},\vec{X},I)=(2\pi)^{-(n+N)/2}/Z\cdot\left(\frac{2\pi}{A}\right)^{1/2}\cdot\int d\gamma\gamma^N$$
$$\cdot\exp\left\{-\frac{1}{2}B(\gamma)\right\} \quad (18)$$

with $B(\gamma)$ as given in Eq. (17b). Note that $B(\gamma)$ is again a quadratic form in γ. We can therefore proceed as before in integrating over γ and replace $B(\gamma)$ by

$$B(\gamma)=Q(\gamma-\gamma_0)^2+R, \quad (19)$$

with

$$R=d^Td-\frac{(\vec{d}^T\vec{f})^2}{(\vec{f}^T\vec{f}+\vec{F}^T\vec{F}\sin^2\vartheta)}-\frac{\vec{f}^T\vec{f}+\vec{F}^T\vec{F}}{g^2\vec{D}^T\vec{D}(\vec{f}^T\vec{f}+\vec{F}^T\vec{F}\sin^2\vartheta)}$$
$$+\frac{2\vec{d}^T\vec{f}\vec{D}^T\vec{F}}{g\vec{D}^T\vec{D}(\vec{f}^T\vec{f}+\vec{F}^T\vec{F}\sin^2\vartheta)}, \quad (20a)$$

$$\cos\vartheta=\vec{D}^T\vec{F}/(\vec{D}^T\vec{D}\cdot\vec{F}^T\vec{F})^{1/2}, \quad (20b)$$

$$Q=\frac{\vec{D}^T\vec{D}(\vec{f}^T\vec{f}+\vec{F}^T\vec{F}\sin^2\vartheta)}{\vec{f}^T\vec{f}+\vec{F}^T\vec{F}}, \quad (20c)$$

$$\gamma_0=\frac{\vec{d}^T\vec{f}\cdot\vec{D}^T\vec{F}-(\vec{f}^T\vec{f}+\vec{F}^T\vec{F})/g}{\vec{D}^T\vec{D}(\vec{f}^T\vec{f}+\vec{F}^T\vec{F}\sin^2\vartheta)}. \quad (20d)$$

Due to the factor γ^N in Eq. (18), marginalization over γ is slightly more complicated than c marginalization. We proceed by expressing γ^N as a polynomial in $(\gamma-\gamma_0)=u$ employing the Taylor expansion

$$\gamma^N=\gamma_0^N\left\{1+N\frac{u}{\gamma_0}+\frac{N(N-1)}{2!}\left(\frac{u}{\gamma_0}\right)^2+\ldots\right\}. \quad (21)$$

The odd terms in u in the expansion Eq. (21) do not contribute to the integral Eq. (18). For the even terms it is easy to derive a recurrence relation by partial integration

$$I_{K+2}=\frac{K+1}{Q}I_K,\quad I_K=\int_{-\infty}^{\infty}z^K\exp\left\{-\frac{1}{2}Qz^2\right\}dz,$$
$$I_0=\left(\frac{2\pi}{Q}\right)^{1/2}. \quad (22)$$

Back substitution of Eq. (22) into Eqs. (21) and (18) finally yields

$$P(\vec{\lambda}|\vec{d},\vec{x},\vec{D},\vec{X},I)=(2\pi)^{-[(n+N)/2]}/Z\cdot\left(\frac{2\pi}{A}\right)^{1/2}\left(\frac{2\pi}{Q}\right)^{1/2}$$
$$\cdot\exp\left\{-\frac{1}{2}R(\vec{\lambda})\right\},$$
$$\gamma_0^N\left\{1+\frac{N(N-1)}{2!}\frac{1}{Q\gamma_0^2}\right.$$
$$\left.+\frac{N(N-1)(N-2)(N-3)}{4!}\frac{1\cdot3}{Q^2\gamma_0^4}+\ldots\right\} \quad (23)$$

and the series in brackets is an Nth or $(N-1)$st order polynomial with only even terms in $\gamma_0\sqrt{Q}$. This completes γ marginalization and the calculation of the marginal posterior $P(\vec{\lambda}|\vec{d},\vec{x},\vec{D},\vec{X},I)$ from which we may now calculate moments of individual components of $\vec{\lambda}$. We have left $\vec{\lambda}$ unspecified so far. The previous analysis applies therefore to any problem with the particular c and γ dependence of Eqs. (10), and (11).

5. Cross Section Model Functions

The choice of cross section model functions will now specify the so far undetermined parameter set $\vec{\lambda}$. Any set of experimental data covers only a finite range of energies. Since the evaluation of rate coefficients at elevated temperatures requires the knowledge of the associated cross sections at even higher energies we base our choice on the Born–Bethe expressions[8] for inelastic collisions. These read

$$\sigma(E)\sim\ln(x)/x,\quad x=\frac{E}{E_0} \quad (24a)$$

for optically allowed transitions and

$$\sigma(E)\sim 1/x \tag{24b}$$

for optically forbidden transitions in the high impact energy limit. We generalize these expressions to

$$\varphi(E)=\frac{(x-1)^{\varepsilon}\ln(x)}{a+(x\ \ 1)^{\varepsilon}\cdot x},\quad \sigma(E)=c\cdot\varphi(E) \tag{25a}$$

for optically allowed transitions and

$$\varphi(E)=\frac{(x-1)^{\varepsilon}}{a+(x-1)^{\varepsilon}\cdot x} \tag{25b}$$

for optically forbidden transitions. Model functions Eqs. (25a) and (25b) contain two adjustable nonlinear parameters a and ε which in terms of our previous notation constitute the components of the vector $\vec{\lambda}=(a,\varepsilon)$. We recall, that the preceding analysis has already accounted for the linear parameter c. The beauty of the particular choice Eq. (25) as a fit function for ionization cross section is that it is asymptotically correct. Moreover, our model functions for the cross sections are positive definite at all energies above threshold. This property cannot in general be guaranteed for a polynomial fit. This completes the outline of how to estimate parameters from the given set of data $(\vec{d},\vec{x},\vec{D},\vec{X})$.

6. Cross Section Estimates

Having obtained parameter estimates with error bars from the posterior probability density Eq. (23) and in much the same way for c from a marginal posterior probability density in c $P(c|\vec{d},\vec{x},\vec{D},\vec{X},I)$ one might expect that the job is done. This is not the case. Using the model function with the posterior values of the parameters amounts to estimating $\langle x^2\rangle$ from $\langle x\rangle$ as $\langle x^2\rangle=\langle x\rangle^2$. It is well known, however, that the difference $\langle x^2\rangle-\langle x\rangle^2$ is equal to the variance of x and the above equality holds if and only if the function from which $\langle x\rangle$ is determined is infinitely sharp. We shall now exploit to what extent this assumption holds for the data which will be used in this paper. Quite generally we want to estimate the value of the ionization cross section σ at an arbitrarily given energy E on the basis of the given data set $(\vec{d},\vec{x},\vec{D},\vec{X})$. We return once more to Bayesian calculus to compute $\langle\sigma\rangle$ from a distribution $P(\sigma|E,\vec{d},\vec{x},\vec{D},\vec{X},I)$. Employing the marginalization and product rules this density may be expressed as

$$P(\sigma|E,\vec{d},\vec{x},\vec{D},\vec{X},I)=\int d\vec{\lambda}\,d\gamma\,dc\,P(c,\gamma,\vec{\lambda}|\vec{d},\vec{x},\vec{D},\vec{X},I)\cdot P(\sigma|c,\vec{\lambda},E,I). \tag{26}$$

The first factor of the integrand is already known from Eqs. (12) and (13). The second factor is the probability density for the cross section σ given c, $\vec{\lambda}$, E. Since the conditions c, $\vec{\lambda}$, $x=E/E_0$ specify the model function for the cross section via Eq. (25) uniquely this is

$$P(\sigma|c,\vec{\lambda},E,I)=\delta(\sigma-c\,\varphi(E,\vec{\lambda})). \tag{27}$$

The resulting distribution Eq. (26) is of scaring complexity. However, this is of minor concern if we content ourselves with expectation values of σ^k from Eq. (26). $k=1$ is of course the cross section mean and $k=2$ is needed to calculate the confidence range $\langle\Delta\sigma^2\rangle=\langle\sigma^2\rangle-\langle\sigma\rangle^2$. Performing the integration over σ first removes the nasty delta function and yields

$$\langle\sigma^k\rangle=\int d\vec{\lambda}\,d\gamma\,dc\{c\cdot\varphi(E,\vec{\lambda})\}^k P(c,\gamma,\vec{\lambda}|\vec{d},\vec{x},\vec{D},\vec{X},I)/Z. \tag{28}$$

The $\vec{\lambda}$ integration in Eq. (28) can be accomplished by elementary numerical integration if $\vec{\lambda}$ is of low dimension. For higher dimensions Markov chain Monte Carlo methods are preferred, in which case $P(c,\gamma,\vec{\lambda}|\vec{d},\vec{x},\vec{D},\vec{X},I)/Z$ constitutes the sampling density and the value of Z does not need to be known! Equation (28) is then the full Bayesian answer to the problem of designing a cross section function on the basis of two different discordant data sets.

7. Cross Section Results

In this section we shall present results of the theory outlined above for the direct and dissociative electron impact ionization cross sections of methane and hydrogen. For methane we have used the same data as Erhardt and Langer in their earlier evaluation namely the experiments of Chatham *et al.*[9] and Adamczyk *et al.*[10] We have chosen the Chatham data $(\vec{d})$ to be correctly calibrated. This choice is of course open to discussion and rests mainly on the fact that Cahatham's data are more recent and were taken in view of the existing Adamczyk data. Both measurements were calibrated against the ionization of argon as a secondary standard. We have chosen g, the prior expectation value of the calibration factor γ of the Adamczyk data, to be equal to $g=1$. A minor problem arises with the Chatham data since the important cross section regions near threshold are presented in their paper[9] as continuous lines. For the present purpose, this representation had to be digitized. The total number of data from Chatham's measurement, that is the cross sections tabulated at selected energies plus the digitized near threshold cross sections, was chosen to be equal to the number of data in Adamczyk's measurement. We have thus, as far as the cross section shape is concerned, attributed equal weight to the two sets of data. Finally we had to assign errors to the data. This is particularly debatable since no error margins are given in either of the source papers.[9,10] A very conservative estimate, which we inferred from the scatter of successive points in the measurements, is 7% of the cross section maximum at all energies in both data sets. This includes digitizing errors!

Figure 2 displays the $CH_4\rightarrow CH_4^+$ experimental data together with the full Bayesian predictive cross section Eq. (28) as a full line and the associated error as the shaded area around this line. Note that the error in the Bayesian estimate is considerably smaller than the 7% (max) attributed to the

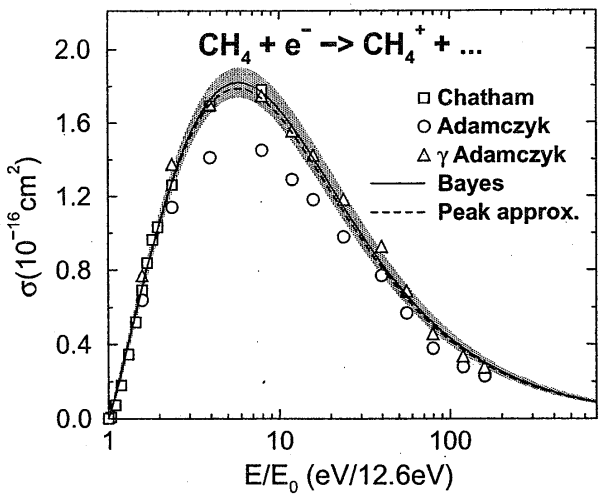

FIG. 2. The full Bayesian result for the cross section fit to the data of Fig. 1 is shown by the solid line. The gray shaded area is the confidence range. From the analysis a correction factor for the data of Adamczyk *et al.* (Ref. 10) may be found to give $\gamma = 1.2$. The open triangles show the data multiplied by this factor. The peak approximation (dashed line) represents the result well within the confidence range.

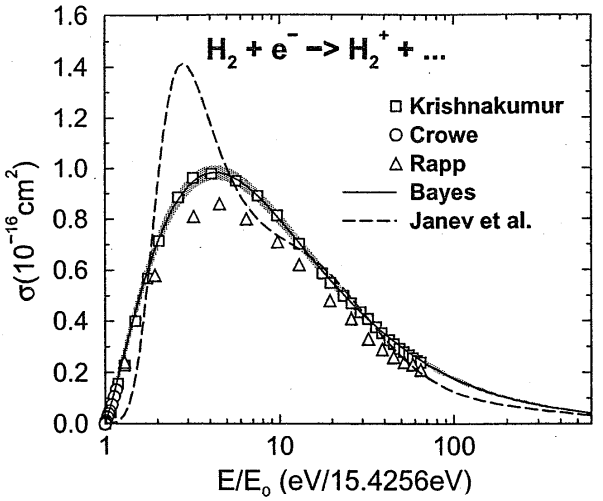

FIG. 3. Cross section for the electron impact ionization of hydrogen. Three different data sets are the basis for the Bayesian result (solid line). The dashed line is a polynomial fit by Janev *et al.* (Ref. 2). Not shown is the peak approximation, since it is hardly distinguishable from the result of the full approach.

experimental data. Moreover, the confidence range is energy dependent unlike the error assignment to the data. This desirable behavior is due to the fact that the model functions Eqs. (25a) and (25b) contain only three parameters, which are estimated on the basis of about ten times as much data. This results in very robust and precise cross section predictions.

In spite of the beauty of the full Bayesian solution, we are well aware that the procedure to obtain the posterior cross section estimates Eq. (28) is by far too involved to be used routinely. It is therefore instructive to evaluate the model function Eqs. (25a) and (25b) with the posterior expectation values of the parameters c, a, ε, and compare it to the full Bayesian solution. Conceptually there is a big difference between the two cases, in fact as large as between $\langle x^2 \rangle$ and $\langle x \rangle^2$ for some distribution $f(x)$. Numerically, however, the two cannot be distinguished within linewidth. Thus the continuous curve in Fig. 2 also represents the model function Eq. (25a) evaluated with parameters fixed at their posterior expectation values. This is due to the precision of the input data and good news for the calculation of rate coefficients (see Sec. 8).

The surprisingly good performance of the model function with parameters fixed to their posterior expectation values has tempted us to try for an even cruder approximation. We assume well-conditioned data such that everywhere, where we meet an exponential possibly multiplied by a slower varying function such as a polynomial, the exponential is replaced by a Dirac δ function positioned at the maximum of the exponential. The only computational effort required in this approximation is to find the maximum of R in Eq. (20a) as a function of the nonlinear parameters $\vec{\lambda}$. Having determined $\vec{\lambda}_{\max}$ by some appropriate search algorithm we find γ_0 from Eq. (19) and finally ϵ_0 from Eq. (17c). This bypass to the full Bayesian calculation produces surprisingly good results in the case of precise data as shown in Fig. 2 by the dashed line. The penalty to this route is that we do not obtain an error margin.

The analysis which we have developed can of course easily be extended to include more than two data sets in the determination of the cross section. A representative case where data from three different experiments were used is the ionization of H_2 into H_2^+. Again we have assumed that the most recent data by Krishnakumar and Srivastava[11] are correctly calibrated and that the earlier data from the work of Rapp *et al.*[12] and Crowe and McConkey[13] need adjustment with prior factor $g = 1$. The experimental data from the three sources are shown in Fig. 3 together with the full Bayesian predictive cross section estimate Eq. (28) as a continuous line. The peak (delta function) approximation is not shown, since it is not distinguishable from the full approach. Also included in Fig. 3 is the polynomial fit of Janev *et al.*[2] as the dotted curve. Its poor reproduction of the experimental data should be taken as a warning against the blind use of their collection of fits at least in case of low temperature plasmas.

The results of this section are summarized in Table 1 which provides posterior estimates for the parameters to be used in model functions Eqs. (25a) and (25b) for all direct and dissociative electron impact ionization cross sections of methane and hydrogen. From the fit result of the respective model it was obvious which model was the appropriate one (shown in the second column) and a Bayesian model comparison between the two choices was not necessary.

8. Rate Coefficients

For an efficient and fast calculation of rate coefficients it is essential that the cross section formulas Eqs. (25a) and (25b)

TABLE 1. Posterior estimates for the model parameters in Eqs. (25a) and (25b).

Channel	Model	c	a	ε	E_0/eV
$CH_4+e^-\rightarrow CH_4^+ + \ldots$	(25a)	9.7	4.4	0.14	12.6
$CH_4+e^-\rightarrow CH_3^+ + \ldots$	(25a)	6.7	3.5	0.20	14.3
$CH_4+e^-\rightarrow CH_2^+ + \ldots$	(25a)	0.93	4.0	1.0	15.1
$CH_4+e^-\rightarrow CH^+ + \ldots$	(25a)	0.31	0.67	1.7	22.2
$CH_4+e^-\rightarrow C^+ + \ldots$	(25b)	0.22	8.7	1.2	25.0
$CH_4+e^-\rightarrow H^+ + \ldots$	(25b)	1.3	22	1.2	18.1
$H_2+e^-\rightarrow H_2^+ + \ldots$	(25a)	3.9	1.8	0.18	15.43
$H_2+e^-\rightarrow H^+ + \ldots$	(25b)	0.91	34	1.5	18.06

when evaluated with the posterior expectation values of the parameters coincide within the linewidth with the result of the full Bayesian calculation Eq. (28). We may for this very reason obtain rate coefficients $\langle \sigma v \rangle$ from Eqs. (25a) and (25b) as

$$\langle \sigma v \rangle = \int_0^\infty v\sigma(E) f(E) dE. \tag{29}$$

For $f(E)$ we choose here a Maxwellian distribution for illustration purposes

$$f(E) = \frac{2}{\pi}\left(\frac{1}{k_B T}\right)^{3/2} \exp\{-E/k_B T\} E^{1/2} \tag{30}$$

We introduce $x = E/E_0$ as a new integration variable and define $\alpha = E_0/k_B T$, $\beta = (2E_0/m)^{1/2}$. Substitution into Eq. (29), noting that $\sigma(E) = 0$ for $E < E_0$, yields

$$\langle \sigma v \rangle = \frac{2}{\sqrt{\pi}} \alpha^{3/2} \beta \int_1^\infty \sigma(x) x \cdot e^{-\alpha x} dx. \tag{31}$$

We shall neither provide tables of Maxwellian averaged rate coefficients nor analytic fits but rather describe their efficient calculation. It can easily be accommodated to a distribution function other than a Maxwellian. Such a need arises if we treat the ill posed inversion problem of deriving the electron energy distribution function from measured spectral line intensities.[14] It is useful to transform the range of integration in Eq. (31) $[1,\infty]$ into $[0,1]$ by introducing $z = x^{-\rho}$

$$\langle \sigma v \rangle = \frac{2}{\sqrt{\pi}} \frac{\alpha^{3/2}\beta}{\rho} \int_0^1 \frac{dz}{z} x^2 \sigma(x) e^{-\alpha x}. \tag{32}$$

The exponent ρ is still arbitrary and can be chosen in a certain sense of optimization. We fix ρ by the requirement that the integrand in Eq. (32) shall peak at $z = 1/2$. The hope is that it is then most "uniformly" distributed in $0 \leq z \leq 1$. The resulting value of ρ is of course temperature (α) dependent and depends in principle also on the particular cross section which is to be averaged. The latter dependence is weak and we have neglected it. This is in fact no approximation, since the choice of ρ influences only the efficiency of executing the numerical integration Eq. (32). We have used the

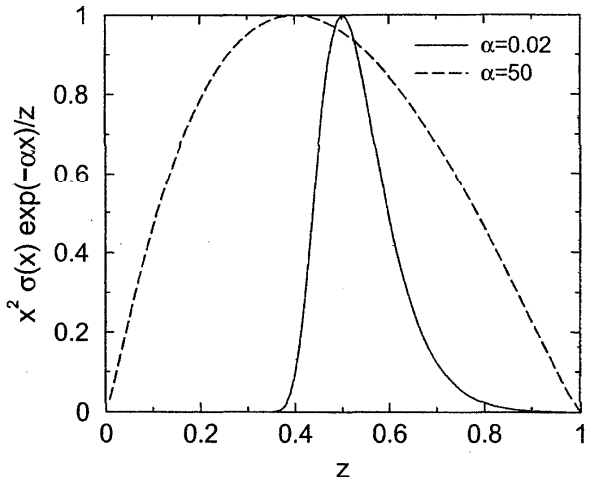

FIG. 4. The integrand of Eq. (32) for $\alpha = 0.02$ (solid line) and $\alpha = 50$ (dashed line). The latter corresponds to the case of low temperature plasmas. It covers the whole range and therefore facilitates numerical integration.

$CH_4 \rightarrow CH_4^+$ cross section as a representative to estimate the optimum transformation exponent ρ. Its temperature dependence is very approximately given by

$$\ln(1/\rho) = 0.519\,30 - 0.630\,72 \ln(\alpha) - 0.076\,07 (\ln(\alpha))^2 \tag{33}$$

in the range $0.02 \leq \alpha \leq 50$. Figure 4 displays the transformed integrand in the range $0 \leq z \leq 1$ for the limits of the validity of Eq. (33) and it is obvious that numerical integration employing the trapezoidal rule recursively will be fast. The authors are ready to supply the subroutine for calculating rate coefficients on request. The same routine also provides temperature derivatives of the rate coefficients from

$$\frac{d}{dT}\langle \sigma v \rangle = \frac{2}{\sqrt{\pi}} \frac{\alpha^{3/2}\beta}{\rho} \int_0^1 \frac{dz}{z} \frac{(\alpha x - 3/2)}{T} x^2 \sigma(x) e^{-\alpha x}. \tag{34}$$

In concluding this section it is worth pointing out that the calculation of the rate coefficient should be very accurate since the presented cross section formulas are asymptotically exact and we therefore do not encounter a truncation problem.

9. Summary

Analytic fit formulas for the determination of cross sections for electron impact ionization of methane and hydrogen were presented. To the knowledge of the authors this work reveals for the first time fit functions for the partial cross sections of CH_2^+, CH^+, C^+, and H^+. Especially for low electron temperatures the new formulas give a better agreement with the experimental data in qualitative and quantitative respect as previous work by, e.g., Erhardt and Langer. While the latter used higher order polynomials for the fit formula, we chose to give the model function a physical

meaning and employed a generalized form of the Born–Bethe expressions for inelastic collisions. This approach already possesses the asymptotically correct form. Our cross section model function contains only two shape parameters and one scale parameter. The data to which this function was fitted came from different sources, covered different energy ranges, and suffered from discordant calibration. In order to cope with this problem the most recent data set was taken as correct on the absolute scale, while the others were equipped with scale factors. Then Bayesian probability theory was employed to evaluate the unknown scale factors, the linear coefficient, and the model parameters, together with the sought for cross section itself. Though the full Bayesian approach has the additional advantage of providing the result with a confidence range, it may be by far too costly to be incorporated in daily use. Therefore two approximations were presented. The first one consists of bypassing the otherwise necessary numerical calculation of the cross section by just using the posterior expectation values of the model parameters and inserting them into the model function. No visible difference with the full approach was found. In the second short cut the model parameters are obtained by finding the maximum in an approximative formula for the generic problem of discordant data sets. Again the result was well within the confidence region of the full Bayesian calculation. Unfortunately, this procedure is not capable of providing an error bar estimation. Both approximations perform better, the higher the quality and number of the data is, a fact which one should bear in mind when moving to other data sets. Finally, a procedure to obtain rate coefficients from the cross sections was presented.

10. References

[1] A. B. Ehrhardt and W. D. Langer, Report PPPL-2477, Princeton (1987).
[2] R. K. Janev, W. D. Langer, K. Evans, Jr., and D. E. Post, Jr., *Elementary Processes in Hydrogen-Helium Plasmas* (Springer, New York, 1987).
[3] P. C. Gregory and T. J. Loredo, Astrophys. J. **398**, 146 (1992).
[4] G. Larry Bretthorst, *Bayesian Spectrum Analysis and Parameter Estimation* (Springer, Berlin, 1988).
[5] *Kendall's Advanced Theory of Statistics, Bayesian Inference*, edited by A. O'Hagan (Wiley, New York, 1994).
[6] D. S. Sivia, *Data Analysis* (Oxford University Press, Oxford, 1996).
[7] J. N. Kapur and H. K. Kesavan, *Entropy Optimization Principles with Applications* (Academic, San Diego, CA, 1992).
[8] H. Bethe, Ann. Phys. (Leipzig) **5**, 325 (1930).
[9] H. Chatham, D. Hils, R. Robertson, and A. Gallagher, J. Chem. Phys. **81**, 1770 (1984).
[10] B. Adamczyk, A. J. H. Boerboom, B. L. Schram, and J. Kistemaker, J. Chem. Phys. **44**, 4640 (1966).
[11] E. Krishnakumar and S. K. Srivastava, J. Phys. B **27**, L251 (1994).
[12] D. Rapp, P. Englaender-Goulden, and D. Briglia, J. Chem. Phys. **42**, 4081 (1965).
[13] A. Crowe and J. W. McConkey, J. Phys. B **6**, 2088 (1973).
[14] R. Fischer, W. Jacob, W. von der Linden, and V. Dose, "Bayesian Reconstruction of Electron Energy Distributions from Emission Line Intensities," in *Maximum Entropy and Bayesian Methods,* edited by W. von der Linden *et al.* (Kluwer Academic, Dordrecht, 1999).